

APPLICATION FOR A UNITED STATES PATENT

UNITED STATES PATENT AND TRADEMARK OFFICE
(MBHB Case No. 01-471; 3Com Case No. 3691.CS.US.P)

5 Title: **Distributed MPLS Architecture**

10 Inventors: Satish Amara, a citizen of the India and a resident of Mount Prospect,
IL;

Shaji Radhakrishnan, a citizen of the India and a resident of Mount
Prospect, IL;

15 Rajesh Ramankutty, a citizen of the India and a resident of
Schaumburg, IL;

20 Sanil Kumar Puthiyandyil, a citizen of the India and a resident of
Lawrence, MA; and

Boby Joseph, a citizen of India and a resident of Mount Prospect, IL.

25

30 Assignee: 3Com Corporation
5400 Bayfront Plaza
Santa Clara, CA 95052

FIELD OF THE INVENTION

This present invention relates to switching information in a network. More specifically, it relates to a system and method for achieving distributed MPLS and packet switching using L2TP as a control mechanism.

BACKGROUND OF THE INVENTION

Multiple Protocol Label Switching (MPLS) networks use a switching technique whereby packets may be routed across a network. The packets transmitted across the MPLS network may take a variety of forms and may include a label. The label may be a fixed value, for example, an integer. The labels may be used to indicate the destination of the packet.

The MPLS network may include a plurality of nodes. The nodes may include Label Edge Routers (LERs) where information enters the network ("ingress nodes") and where information leaves the network ("egress nodes"). The LER may add a label to the head of the packet to indicate the destination of the packet. The LERs may ignore other information in the packet, for example, Internet protocol (IP) addresses and ATM VCI/VPI information.

The LER may be used in a MPLS network as the boundary between Layer 3 forwarding and MPLS forwarding. The LER may include functionality to add a label to an unlabeled packet ("an ingress LER") and remove labels from the packet ("an egress LER").

Label Switching Routers (“LSRs”) may be used to route the packets between LERs. The LSRs may examine the label in a packet to determine the destination of the packet. In one example, the label may indicate an index in a table (stored in the switching node) and may be used to determine the outgoing link to which the packet
5 may be forwarded. The table may be stored in a memory at the switching node, for example.

The LSRs may assign a new label and forward the packet on the link. Each label may have significance only locally. In other words, the packets may be forwarded hop-by-hop across the MPLS network. The label may indicate each hop
10 rather than the entire end-to-end path from the source to the destination.

SUMMARY OF THE INVENTION

The system and method of the present invention advantageously provides for
5 the distributed processing of labeled packets in a device. For example, a first type of
packet may be processed by an ingress module and a second type of packet may be
processed by a route server module.

In one example of the present invention, a system for processing packets of
information includes an ingress module, which is coupled to a route server module.

10 The ingress module may receive a plurality of packets of information from a
first network and may determine the type of each of the plurality of packets. The
route server module may send a distributed processing request to the ingress module.

The ingress module may receive the distributed processing request and,
responsively, may perform a first set of processing operations on selected ones of the
15 plurality of packets. The ingress module may receive the FTN and NHLFE tables
from router server. The selected ones of the plurality of packets may be of a first
type. The ingress module may forward others of the plurality of packets of
information to the route server module. Each of the others of the plurality of packets
may be of a type distinct from the first type.

20 The route server module may receive the others of the plurality of packets of
information and performs a second set of processing operations on the others of the
plurality of packets of information.

The first set of processing operations may include forwarding the selected
ones of the plurality of packets of information to an egress module. The second set of

processing operations includes establishing a connection with an entity on the Internet. The first type of packet may be a data type.

- The system may further include an egress module, and the egress module may be coupled to the ingress module. The egress module may receive the others of the
- 5 plurality of packets and route the packets to the Internet.

These as well as other aspects and advantages of the present invention will become more apparent to those of ordinary skill in the art by reading the following detailed description, with reference to the accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Preferred embodiments of the present inventions are described with reference to the following drawings, wherein:

Figures 1a and 1b are diagrams illustrating a preferred embodiment of the system for distributed MPLS processing in accordance with the present invention;

Figure 2 is a call flow diagram illustrating distributed MPLS processing in accordance with a preferred embodiment of the present invention;

Figure 3 is a diagram illustrating a distributed switching request in accordance with a preferred embodiment of the present invention; and

Figure 4 shows a diagram showing a device for implementing distributed MPLS processing in accordance with a preferred embodiment of the present invention.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

Referring now to Figure 1a, a system includes a user device 102, a label edge router (LER) 104, a plurality of label switch routers (LSRs) 106, a LER 108, and a user device 110. The user device 102 is coupled to the LER 104. The LER 104 is
5 coupled to the LSRs 106. The LSRs 106 are coupled to the LER 108. The LER 108 is coupled to the user device 110.

The user devices 102 and 110 may be any type of device used to transmit and/or receive information. In one example, the user device may be a personal computer. Other types of user devices are possible.

10 The functions of the LERs 104 and 108 may be implemented by a processor executing computer instructions stored in a memory. The LERs 104 and 108 may include an ingress module, egress module, and route server module, as described elsewhere in this specification. The LERs 104 and 108 may receive packets from the user devices and insert a label into these packets and forward the packets to the LSRs
15 106. The LERs 104 and 108 may also perform distributed switching, which is also described elsewhere in this specification.

The functions of the LSRs 106 may be implemented by processors executing computer instructions stored in a memory. The LSRs 106 may include an ingress module, egress module, and route server module, as described elsewhere in this
20 specification. The LSRs 106 may receive a packet having a label and route the packet to the next destination. In the routing process, the LSRs 106 may replace the current label with a new label. The new label may signify the destination of the

packet. The LSRs 106 may also perform distributed switching, which is described elsewhere in this specification.

The LERs 104 and 108 may determine a forwarding equivalence class ("FEC") for the incoming packets that, based on the assigned FEC, are forwarded in the same manner (e.g., over the same path, with the same forwarding treatment). The assignment of a particular FEC to a particular label may be done once, as the packet enters the network, and the FEC to which the packet is assigned is encoded as a label. When the packet is forwarded to its next hop, the label may be sent along with it, i.e., the packets may be labeled before they are forwarded. At subsequent hops, there is no further analysis of the packet's network layer header. Rather, the label is used as an index into a table that specifies the next hop and a new label. At subsequent hops, the LSRs 106 may use the information from the packet to determine the outgoing link and a new label for the outgoing link. The LSRs 106 then may swap the label in the MPLS header with a new label, and forward the packet.

Each LER 104 and 108 or LSR 106 may negotiate a label for each FEC with its neighbors along the path. Information on the topology of the network may be maintained by one or more routing protocols such as an open shortest path first ("OSPF"), a routing information protocol ("RIP"), or a border gateway protocol ("BGP"), for example. For each route or aggregation of routes, a neighbor router may assign a label, and this information may be distributed to neighboring LERs 104 and 108 or LSRs 106 using a label distribution protocol (LDP) or can be piggybacked on BGP route updates (RFC 3107, Carrying label information). For example, the system

may use the RFC-3036 protocol developed by the Internet engineering task force ("IETF").

Referring now to Figure 1b, a device 150 includes a route server module 152, an ingress module 154, an egress module 156, a network 158, and a network 160.

5 The route server module 152 is coupled to the ingress module 154. The ingress module 154 is coupled to the egress module 156 and the network 160. The egress module 156 is coupled to the network 158.

09443971.083101
The functions of the route server module 152 may be implemented by a processor executing instructions stored in a memory. The route server module 152 may receive and route IP data packets, before the sending of a distributed switching message to the ingress module 154. The route server module 152 may send a message to the ingress module 154 asking the ingress module 154 to process all data packets received from the PSTN. The route server module 152 may also process all control messages and IP packets having local end points. The route server module 152 may perform other functions as well. The route server module may send FTN and NHLE entries to the ingress module for label swapping.

The route server module 152 may send the message to the ingress module 154 asking the ingress module to process all data packets upon the occurrence of a predetermined condition. For example, at the time the route server module 152 completes the PPP negotiation process, this message may be generated.

The functions of the ingress module 154 may also be implemented by a processor executing instructions stored in a memory. The ingress module 154 may receive IP packets from the network 160 and determine the type of packet. For

example, the ingress module 154 may determine whether the IP packet is a control packet, a data packet, or any packet destined for a local connection. Based upon this determination, the ingress module 154 may route the packet to the egress module 156, route server module 152, or perform further processing itself.

5 The ingress module 154 may also perform distributed forwarding. The ingress module 154 may, for example, route IP data messages to the network 158 after receiving a distributed switching request.

 In addition, the ingress module 154 may receive messages from the network 158, process the messages, and forward the messages to a destination. In one
10 example, an IP packet may be received by the ingress module 154 from the network 158. The ingress module 154 may determine the destination of the IP packet, encapsulate the packet with an PPP header, and forward the encapsulated packet to a destination on the network 158.

 The functions of the egress module 156 may also be implemented by a
15 processor executing instructions stored in a memory.

 The network 158 may be any network capable of transporting any type of information. For example, the network may be the Internet and transport IP packets. In addition, the network 158 may be a combination of networks. Other examples of networks are possible.

20 The network 160 may be any network capable of transporting any type of information. For example, the network may be a PSTN and transmit information according to the point-to-point protocol (PPP).

In one example of the operation of the system of Figure 1b, the ingress module 154 initially tunnels all PPP packets coming from the network 160 to the route server module 152. The L2TP protocol is used to tunnel PPP packets from ingress to route server module. In this example, ingress router acts as LAC and router server as LNS.

5 The route server module 152 processes the packets. For instance, the route server module 152 may perform MPLS negotiation, PPP negotiation, and determine IP network for the link with the network 160.

The route server module 152 may send a control packet, for example, and L2TP control packet, to the LAC within the ingress module 154. The control packet

10 may request that distributed switching may take place. The control packet may also contain FTN and NHLE tables for label swapping. The ingress module 154 may send a response message, for example, a response packet acknowledging the receipt of the control packet. The control packet may cause the ingress module 154 to halt the forwarding data packets to the route server module 152, and, instead, keep the packets

15 for further processing. The ingress module 154 may also receive updated label swapping and forwarding table from the route server module 152.

The ingress module 154 may strip off the PPP header and perform decompression, if needed. The ingress module 154 may then forward the packet to the egress module 156.

20 Incoming packets (from the network 158) may be received at the egress module 156 and forwarded to the ingress module 154. The ingress module 154 may encapsulate the packets with a header and may perform compression, label swap and transmit the packets over a link to the network 160.

The ingress module 154 may also route packets coming from the network 160 destined for PPP local endpoints (indicated by the IP addresses), to be sent to the route server module 152.

Referring now to Figure 2, a method of distributed switching is described in reference to a system that includes an ingress module, which is coupled to a route server module. An egress module may be coupled to the ingress module. The ingress module may include a LAC and a distributed forwarding agent, and the route server module may include an LNS. The ingress module may be coupled to a PSTN and the Internet. The route server module may be coupled to the Internet.

At step 202, PPP negotiation packets are sent from an outside source, for example, from a user, to the ingress module. For example, the PPP negotiation packets may be sent to the ingress module.

At step 204, a tunnel is created between the ingress module and the route server module. For example, the tunnel may be established according to the L2TP protocol. Other protocols may also be used.

At step 206, PPP negotiation packets are sent from an outside source, for example, from a user, to the ingress module. For example, the PPP negotiation packets may be sent to the ingress module.

At step 208, a tunnel is created between the LAC and the route server module. For example, the tunnel may be the same tunnel established with the LNS in the route server module according to the L2TP protocol.

At step 210, the LNS in the route server module sends a message to the ingress module to tell the ingress module to distribute the switching of all subsequently received packets.

At step 212, a response message is sent from the LAC in the ingress module to
5 the LNS in the route server module.

From this point, at steps 214, 216, and 217, all PPP encapsulated outgoing data packets from the PSTN network to the Internet will be forwarded to the distributed switching agent in the ingress module. The ingress module will also get updated swapping and forwarding tables from the route server module to support the
10 forwarding. The ingress module may remove the PPP header and give the IP data packets to the distributed forwarding agent in the ingress module. All incoming IP packets reaching the distributed switching module for the PPP link will be given to the ingress module. The ingress module will encapsulate the PPP header and may compress the packet. The ingress module may also perform label swapping and send
15 the packet over the PPP link.

At steps 218 and 220, all IP packets coming from the PPP link destined for PPP local endpoint addresses are sent to the LNS in the route server module. These packets include ICMP, RIP, and other routing protocol packets, for example.

At step 222, PPP control packets coming from the PPP link are received at the
20 ingress module. At step 224, the PPP control packets are tunneled to the LNS in the route server module.

At step 226, MPLS LDP, CRLDP and RSVP-TE packets are received at the ingress module. At step 228, these packets are tunneled to the LNS in the route server module.

Referring now to Figure 3, one example of a distributed forward request message is described. The message may be in the form of an attributed value pair (AVP) 300. The AVP 300 may include a type field 302, a length field 304, and a value field 306. In one example, the type field may be set to "distributed forwarding request," the length field may be set to 2, and the value field may remain empty. Other examples of messages and field values are possible.

Referring now to Figure 4, one example of a system 400 for distributed switching is described. An ingress module 402 includes a LAC 404, a MPLS label switch 406, and a MPLS distributed forwarding agent 408. The functions of any of these elements may be implemented using a processor executing instructions stored in a memory. The ingress module 402 may be coupled to an egress module 419. The egress module 419 may be coupled to the PSTN 420.

The system 400 also includes a route server module 410. The route server module 410 includes an LNS module 412 and a centralized routing module 414. The functions of any of these elements may also be implemented using a processor executing instructions stored in a memory. The system 400 is coupled to a PSTN 420 and the Internet 422. The system 400 may be an LER, LSR, or any other type of device that routes packets or any other type of information.

module) to the packets and switch the packets to a destination. The MPLS label switch 406 may also receive packets from the Internet and forward the packets to the LAC 404 for processing for example, if there is no FTN entry for the packet.

The MPLS distributed forwarding agent 408 may label the packets using the table received from the LAC 404. The MPLS distributed forwarding agent 408 may also receive packets from the egress module 419 and route the packets to the LAC 402.

The LNS 412 may supply label tables to the LAC 404. The LNS 412 may also receive packets from the LAC 404 to be routed to a destination, control packets, negotiation packets, or any other type of packets. The LNS 412 may forward these to the centralized routing module 414.

The centralized routing module 414 supplies transmission rules to the MLPS label switch. The centralized routing module 414 also may route packets (received via the LNS) to a destination on the Internet 422.

In one example of the operation of the system of Figure 4, a control packet may be received by the MPLS label switch 406. The packet may be a PPP negotiation packet and the MPLS label switch 406 may not contain a rule for this type of packet. The MPLS label switch 406 may forward the packet to the LAC 404. The LAC 404 may forward the packet to the LNS 412. The LNS 412 may forward the packet to the centralized routing module 414. The centralized routing module 414 may perform whatever service is required (e.g., PPP negotiation).

After negotiation is completed by the route server module 410 and centralized routing module 414, a MPLS distributed switching packet may be sent from LNS 412

to LAC 404. The MPLS distributed switching packet may inform the LAC 404 to begin performing distributed switching. The LAC 404 may send an acknowledgement packet.

Subsequently, data packets may be received at the MPLS label switch 406 at the ingress module 402. The MPLS label switch 406 may include a filter module, which is coupled to the MPLS label switch and the LAC module. The filter module may contain filter rules and actions to be taken when filter rules are matched. For example, the filter rules can be PPP negotiations, MPLS control packet and actions to be taken is the packets are forwarded to LAC. By default, if there is no matching rule then packets are forwarded to MPLS label switch. This functionality can also be integrated in MPLS label switch.

The ingress module may examine the packets, check the packet type, and determine that the packets are data packets. For example, the packet may have a type field. The algorithm may examine the type field and from the examination determine the type of packet. Alphanumeric characters may be used to indicate the type. Other mechanisms and algorithms may also be used. The MPLS distributed forwarding agent 408 may place a label in the packets. The MPLS label switch 406 may forwards the packet to the Internet 422 via the egress module 419, without involving the route server module 410.

The MPLS label switch 406 at the ingress module 402 may also subsequently receive control or other non-data packets. The ingress module may examine these packets, determine the packets are non-data packets and transmit the packets to the

LNS 412 in the route server module 410. The LNS 412 may route the packets to the centralized routing module 414.

In view of the wide variety of embodiments to which the principles of the present invention can be applied, it should be understood that the illustrated
5 embodiments are exemplary only, and should not be taken as limiting the scope of the present invention. While various elements of the preferred embodiments have been described as being implemented in software, in other embodiments in hardware or firmware implementations may alternatively be used, and vice-versa.

It will be apparent to those of ordinary skill in the art that methods involved in
10 the system and method for a distributed MPLS architecture may be embodied in a computer program product that includes a computer usable medium. For example, such a computer usable medium can include a readable memory device, such as, a hard drive device, a CD-ROM, a DVD-ROM, or a computer diskette, having computer readable program code segments stored thereon. The computer readable
15 medium can also include a communications or transmission medium, such as, a bus or a communications link, either optical, wired, or wireless having program code segments carried thereon as digital or analog data signals.

The claims should not be read as limited to the described order or elements unless stated to that effect. Therefore, all embodiments that come within the scope
20 and spirit of the following claims and equivalents thereto are claimed as the invention.